NatureAtlas.org: Exposing Undergraduate Computer Science Students to Opportunities and Applications in Biodiversity Informatics

Nazli W. Hardy and Christopher R. Hardy

***Abstract*—Biodiversity informatics is a recently emerged application of computer science and information technologies to the quest to better document, track and conserve species of flora and fauna. To this end, undergraduate computer science students should be exposed to the theory and practice of how they can contribute to the development of this field. Since 2007, computer science students at Millersville University of Pennsylvania have worked with biology students and faculty to develop NatureAtlas.org, a Web application to which occurrences of plants, animals, fungi and associated data can be mapped interactively and shared globally on the Web. For over seven years, computer science students have had the satisfaction of seeing their product expand in its use to other universities and non-profit organizations. Their efforts have helped lay the foundation for the generation of valuable data for thousands of species and their occurrences in nature. We hope to expand such educational development opportunities to additional students and collaborations with additional universities.**

***Index Terms*—Biodiversity informatics, undergraduate computer science research, multidisciplinary application of computer science, interactive Web application**

BIODIVERSITY informatics is a relatively new branch of computer science that can be described as the application of computer science and related information technologies to the collection, curation, dissemination, analysis, and interpretation of information regarding biodiversity [1]. Biodiversity informatics now plays a critical role in both the discovery and conservation of species. In the midst of a looming biodiversity crisis characterized by accelerated rates of anthropogenic extinction [2], biodiversity informatics is a critically important application of computer science to which undergraduate university students in both computer science and biology should be given formal exposure.

One important aspect of biodiversity informatics that is easily integrated into undergraduate courses and research projects is the collection of species occurrence records. Occurrence records provide data about where and when plants and animals are known to occur in geographic space, and such data have long been collected through traditional means (i.e., into paper notebooks or spreadsheets) by biology students as part of laboratory field excursions or specimen collection or journaling projects. Without the aid of information technology, however, such student-generated biological data are relegated to the obscure annals of gradebooks and student notebooks. The true potential of such data to inform biologists and conservationists about where species exist and whether they are spreading or in decline and threatened with extinction is not realized since data are not widely if at all accessible on the Web. Since 2007, undergraduate computer science students and faculty at Millersville University of Pennsylvania have worked with biology students and faculty to develop NatureAtlas.org, a Web-based application that has brought biological field excursions and projects at Millersville and other universities into the 21st century.

Fig. 1. The NatureAtlas.org homepage (www.natureatlas.org). The various icons are linked to organism-specific portals for birds, fish, fungi, herps (amphibians and reptiles), invertebrates, mammals, plants and zooplankton.

NatureAtlas.org (www.natureatlas.org; Fig. 1) runs on standard internet browsers and employs a MySQL database and the Google Maps API to allow students to record their observations of plants and animals directly and precisely onto an interactive map and into the database where they are then available on the Web for query, mapping, and download (Figs. 2 and 3). Although primarily used for course projects and student research projects, this site can be used by anyone with Web-access to report new occurrences or to download data about biodiversity.

Since 2007, the NatureAtlas.org project has supported 10 computer science undergraduate students in research

Manuscript received March 26, 2016; revised April 1, 2016. This work was supported in part by a Millersville University of Pennsylvania Faculty Development Grant to C. R. Hardy, as well as a Pennsylvania State System of Higher Education Research Grant to N. W. Hardy and C. R. Hardy.

N. W. Hardy is with the Department of Computer Science and the James C. Parks Herbarium, Millersville University of Pennsylvania, PO Box 1002, Millersville, PA, 17551-0302 USA (phone: 717-871-4312; fax: 717-871-7963; e-mail: nazli.hardy@millersville.edu).

C. R. Hardy is with the Department of Biology and the James C. Parks Herbarium, Millersville University of Pennsylvania, PO Box 1002, Millersville, PA, 17551-0302 USA (phone: 717-871-4317; fax: 717-871-7964; e-mail: christopher.hardy@millersville.edu).

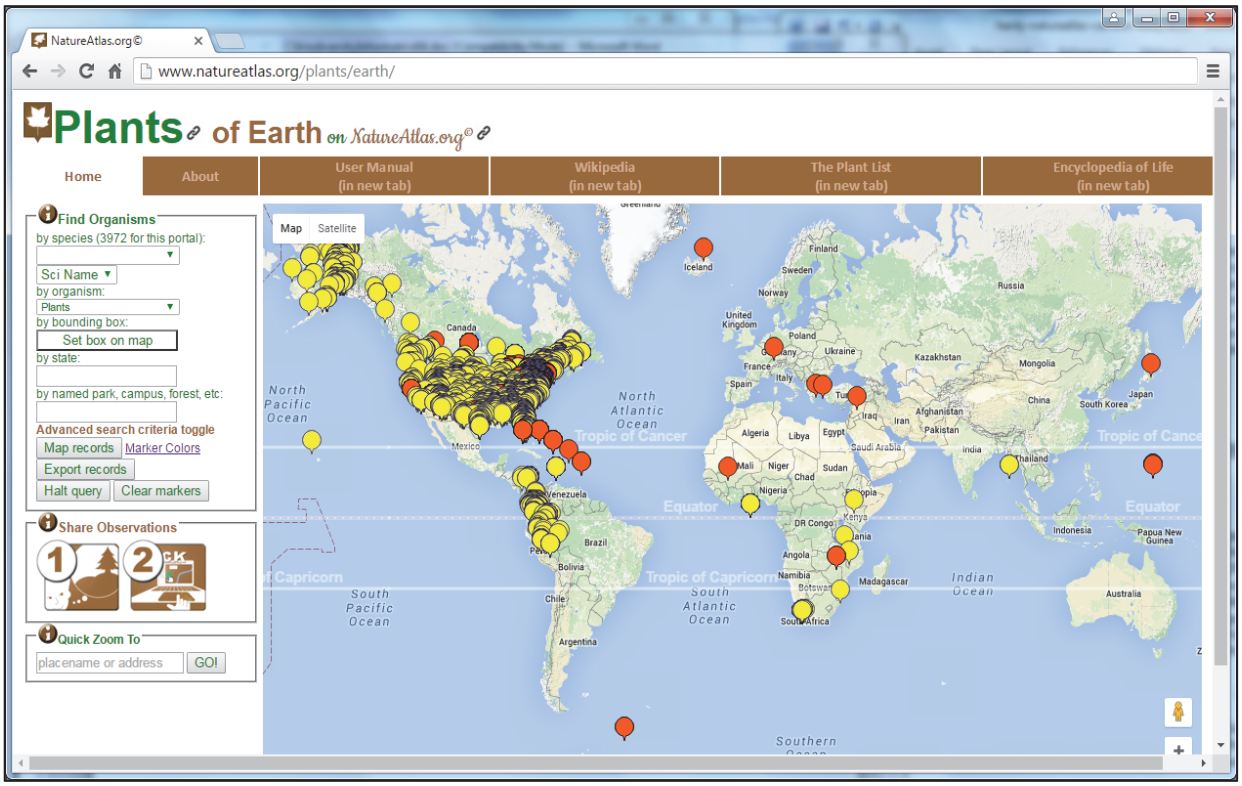

Fig. 2. The *Plants of Earth* portal of NatureAtlas.org (www.natureatlas.org/plants/earth/) provides access to over 18,000 plant occurrence records globally.

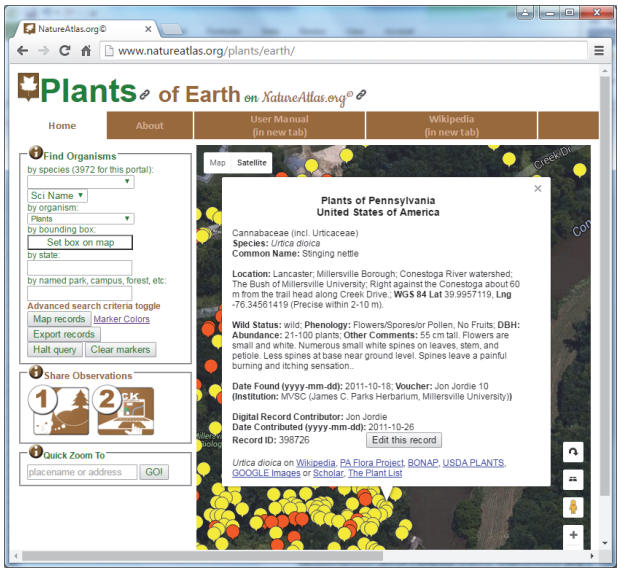

Fig. 3. Each occurrence record's marker can be clicked on for detailed information about the *who*, *what*, *when* and *where* of the record.

projects where they have worked with biology faculty and students, learning how to apply their skills in the creation of an application that enhances the work of their biology counterparts at the university. Computer science students on the project have had the opportunity to work with a large and diverse biology student body [3], [4] and have the satisfaction to have seen use of the Atlas expand considerably since its release in 2008 [5]-[7]. In the seven complete years that NatureAtlas.org has been operational

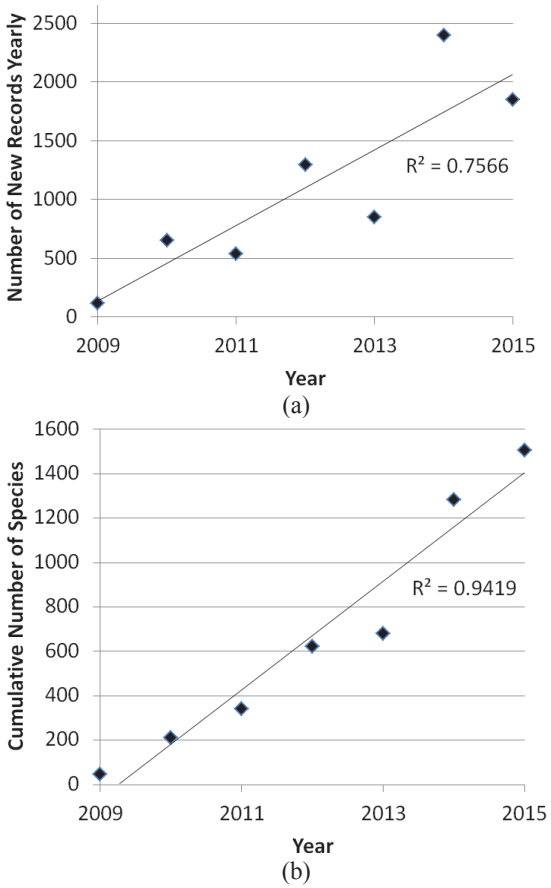

Fig. 4. NatureAtlas.org's rate of growth in wild occurrence records (a) and its accumulation of species reported in the wild (b) have increased steadily in the seven years since its release.

ISBN: 978-988-19253-0-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

(2009-2015), the rate of growth in new records and new species added to the Atlas have increased steadily each year to amount to more than 7,700 new occurrence records in the wild for over 1,500 species of plants, animals, and fungi (Fig. 4). Many more records for cultivated plants and domestic animals also have been generated. NatureAtlas.org has also been integrated into the biology curriculum at five additional universities (Table I) and has been used by students to coordinate the collection of data for independent research projects and for events called BioBlitzes where members of the public and scientific community converge upon a natural area for 24 hours in order to catalog as many species as possible and to foster appreciation for and knowledge about biodiversity.

TABLE I

INSTITUTIONS USING NATUREATLAS.ORG

Institution	NatureAtlas Portal
Climbers Run Nature Preserve (Lancaster County Conservancy)	www.natureatlas.org/climbers/
Fairmont State University	www.natureatlas.org/plants/fairmont/
Hope College	www.natureatlas.org/plants/hope/
Millersville University	www.natureatlas.org (various portals)
Minot State University	www.natureatlas.org/plants/minot/
Rhodes College	www.natureatlas.org/plants/rhodes/
Sinnemahoning State Park (Pennsylvania Department of Conservation and Natural Resources)	www.natureatlas.org/plants/sinnemahoning/
University of Pittsburgh	www.natureatlas.org/plants/pitt-bradford/

In conclusion, biodiversity informatics is an area ripe for the engagement of undergraduate students in collaborative applications of Computer Science that give them exposure to opportunities beyond their traditional scope of study. We welcome further and ongoing collaboration with additional students and faculty at Millersville and other universities.

ACKNOWLEDGMENT

We thank Andrew Bonsall, Mark Olszewski, Matt Snyder, Matt Boaman, Dan Henrich, Brian Houder, Matt Kibler, John Ross, Trevor Scheitrum, and Thomas Murillo for their work as project interns.

REFERENCES

[1] I. N. Sarkar, "Biodiversity informatics: the emergence of a field," *BMC Bioinformatics*, vol. 10, suppl. 14, p. S1., 2009.
[2] D. B. Wake and V. T. Vredenburg, "Are we in the midst of the sixth mass extinction?," *Proceedings of the National Academy of Sciences USA*, vol. 105, suppl. 1, pp. 11466-11473, 2008.
[3] National Science Foundation, "Women, Minorities, and Persons with Disabilities in Science and Engineering," 2015. Available: http://www.nsf.gov/statistics/2015/nsf15311/
[4] National Center for Education Statistics, "Digest of Education Statistics," 2013. Available: http://nces.ed.gov/programs/digest/d13/
[5] C. R. Hardy, N. W. Hardy and N. P. Hartley, "Integrating IT, Web 2.0, and traditional approaches in the plant taxonomy classroom to engage the next generation of professional and citizen biodiversity scientists," *Proceedings of Botany 2012, the Annual International Meeting of the Botanical Society of America, American Association of Plant Taxonomists and International Association of Plant Taxonomists*, July 2012. Available: http://2012.botanyconference.org/engine/search/index.php?func=detail&aid=622
[6] D. Trisel and C. R. Hardy, "The Fairmont State Flora Project on Wiki-Plant-Atlas©," *Proceedings of the 87th Annual Meeting of the West Virginia Academy of Science*, April 2012.
[7] C. R. Hardy and N. W. Hardy, "Use of the undergraduate classroom to advance an emerging paradigm in the study and conservation of biodiversity," *Proceedings of Monocots V, the 5th International Conference on the Comparative Biology of the Monocotyledons*, July 2013. Available: https://www.regonline.com/custImages/320000/329272/July8NYBGMonocotsVAbstractBook.pdf